

昆虫比较线粒体基因组学研究进展*

魏书军^{1**} 陈学新^{2***}

(1. 北京市农林科学院植物保护环境保护研究所 北京 100097; 2. 浙江大学昆虫科学研究所 杭州 310029)

摘 要 动物线粒体基因组因其基因组稳定、基因排列相对保守、普遍为母系遗传、极少发生重组等而被广泛应用于进化与系统发育等研究。目前,昆虫中已有 356 个线粒体基因组序列被测定,代表了 33 个目中的 28 个目。大量比较基因组学研究使得我们对昆虫线粒体基因组的特征与进化方式有了较为清晰的认识。本文对昆虫线粒体基因组的测序进展、基因组的结构特征、碱基组成、控制区的特征、基因重排及其机理、进化速率及其在昆虫系统发育研究中的应用等方面的研究进展进行介绍。

关键词 六足动物, 比较基因组学, 基因重排, 进化速率, 控制区, 系统发育

Progress in research on the comparative mitogenomics of insects

WEI Shu-Jun^{1**} CHEN Xue-Xin^{2***}

(1. Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China; 2. Institute of Insect Sciences, Zhejiang University, Hangzhou 310029, China)

Abstract Animal mitochondrial genomes are widely used in evolutionary and phylogenetic studies because of their stable gene content, relatively conserved gene arrangement, maternal inheritance and infrequent recombination. 356 insect mitochondrial genomes have so far been sequenced from 28 of 33 insect orders. The progressive works on comparative genomics enrich our knowledge of the characteristics and evolutionary patterns of insect mitochondrial genomes. Here we review progress in this field, including features of genome structure, base composition, characteristics of the control region, gene rearrangement and its possible mechanisms, evolutionary rates and the application of insect mitochondrial genomes in phylogenetic research.

Key words Hexapoda, comparative genomics, gene rearrangement, evolutionary rate, control region, phylogeny

线粒体是一种几乎存在于所有真核生物的细胞器,与能量代谢、凋亡、衰老以及疾病相关,是氧化磷酸化的场所。线粒体具有自己的遗传物质,即线粒体基因组(Nass and Nass, 1963)。动物线粒体基因组具有许多共有特征,如基因组稳定、基因排列相对保守、普遍为母系遗传、极少发生重组等(Wolstenholme, 1992),因而被广泛用于进化、系统发育、种群遗传结构、基因漂流、杂交和生物地理学等方面的研究(Wilson *et al.*, 1985)。不同类群的线粒体基因组又表现出特殊的进化方式与特征,因此给基因组的进化研究以及利用线

粒体基因组中的分子标记带来了机遇与挑战。昆虫(六足总纲)作为地球上种类最为繁多、生物学习性最为复杂的生物,一直以来都是线粒体基因组研究的热点类群(Boore, 1999)。近些年,随着测序技术的成熟和线粒体基因通用引物的应用(Simon *et al.*, 1994, 2006),被测定的昆虫线粒体基因组的数量迅速增加,昆虫线粒体基因组特征与进化规律越趋明显,在系统发育研究等方面的应用逐渐成熟。本文对昆虫比较线粒体基因组学的研究进展进行介绍,以期对线粒体基因组的研究和利用提供新的思路。

* 资助项目:国家自然科学基金(31101661; 30970384)、国家自然科学基金创新群体项目(31021003)、国家重点基础研究发展计划(2009CB119004)、北京市科技新星计划(2010B027)、北京市优秀人才培养资助(2010D002020000010)、北京市农林科学院科技创新能力建设专项(KJ CX201104009)。

**E-mail: shujun268@163.com

***通讯作者, E-mail: xxchen@zju.edu.cn

收稿日期: 2011-10-22, 接受日期: 2011-11-08

1 昆虫线粒体基因组测序进展

目前已经测定 356 个昆虫线粒体基因组序列 (GenBank 截止到 2011 年 9 月 6 日), 除纺足目、革翅目、缺翅目、虱目和毛翅目外, 另外 28 个目均有代表性种类的线粒体基因组序列被测定。测定数量较多的目为双翅目、鞘翅目、半翅目、鳞翅目、膜翅目、直翅目和竹节虫目。我国在昆虫线粒体基因组的序列测定中作出了巨大贡献, 测定涉及 14 个目 115 个线粒体基因组。其中直翅目、鳞翅

表 1 世界和我国测定的昆虫线粒体基因组数量比较

Table 1 Comparison on the quantities of the presently sequenced insect mitochondrial genomes world widely and in China

目 Order	拉丁 学名 Scientific name	世界 World	我国 China	我国占 的比例 (%) Proportion of China
双翅目	Diptera	59	8	13.6
鞘翅目	Coleoptera	58	2	3.4
半翅目	Hemiptera	53	34	64.2
鳞翅目	Lepidoptera	40	21	52.5
直翅目	Orthoptera	38	26	68.4
膜翅目	Hymenoptera	32	12	37.5
竹节虫目	Phasmatodea	17	1	5.9
弹尾目	Collembola	9	0	0
蜉蝣目	Ephemeroptera	5	4	80
等翅目	Isoptera	5	0	0
脉翅目	Neuroptera	5	0	0
石蛎目	Archaeognatha	4	1	25
蜻蜓目	Odonata	4	1	25
虱目	Phthiraptera	4	0	0
蜚蠊目	Blattodea	3	2	66.7
双尾目	Diplura	3	0	0
广翅目	Megaloptera	3	1	33.3
捻翅目	Strepsiptera	3	0	0
缨尾目	Thysanura	3	0	0
蛩蠊目	Grylloblattodea	1	0	0
螳螂目	Mantodea	1	0	0
螳螂目	Mantophasmatodea	1	0	0
长翅目	Mecoptera	1	1	100
襀翅目	Plecoptera	1	0	0
原尾目	Protura	1	1	100
啮虫目	Psocoptera	1	0	0

注: 数据来自 GenBank, 截止到 2011 年 9 月 6 日, 包含没有完全测通的线粒体基因组序列。

Data from GenBank by the end of September 6th, 2011, including the nearly complete mitochondrial genomes.

目和半翅目一半以上的线粒体基因组是由我国研究人员测定的 (表 1)。

2 昆虫线粒体基因组的结构与基因组成

2.1 昆虫线粒体基因组的结构

昆虫线粒体基因组为环状闭合双链 DNA, 长度为 16 kb 左右, 通常稳定的包含 13 个蛋白编码基因、22 个转运 RNA (tRNA) 和 2 个核糖体 RNA (rRNA) 基因, 共 37 个基因, 另外还有一段较长的非编码区, 称为控制区或 A + T 富含区 (Boore, 1999)。少数种类会出现同一个基因的多个拷贝或多个控制区, 但这些重复的基因或控制区可能会因为突变而失去功能 (Shao, 2003a; Castro *et al.*, 2006; Cameron *et al.*, 2008)。除虱子中部分种类的线粒体基因组由多个小环组成外 (Shao *et al.*, 2009; Cameron *et al.*, 2011), 其他所有已报道的种类的线粒体基因均位于同一个环状基因组上。昆虫线粒体基因内没有发现内含子, 基因之间的间隔区较小或相邻基因之间有重叠区域 (Boore, 1999)。线粒体基因组的大小主要受非编码区的影响, 目前报道的最大的昆虫线粒体基因组来自黑腹果蝇 *Drosophila melanogaster*, 长 19 517 bp, 其次是半闭弯尾姬蜂 *Diadegma semiclausum*, 长 18 728 bp, 二者均具有较长控制区 (Lewis *et al.*, 1995; Wei *et al.*, 2009)。

2.2 昆虫线粒体基因

昆虫线粒体基因组上的多数基因在同一条链上编码, 该链称为 J 链 (majority strand), 少数基因在另一条链上编码, 称为 N 链 (minority strand) (Simon *et al.*, 1994)。蛋白质编码基因分别为细胞色素 C 氧化酶亚基 I、II、III (*cox1*、*cox2*、*cox3*), 细胞色素 b 脱辅基酶 (*cob*), ATP 合成酶亚基 6 和 8 (*atp6* 和 *atp8*), NADH 脱氢酶亚基 1-6 (*nad1-6*) 和 4L (*nad4l*)。22 个 tRNA 基因中有 18 个 (*trnA*、*trnR*、*trnN*、*trnD*、*trnC*、*trnQ*、*trnE*、*trnG*、*trnH*、*trnI*、*trnK*、*trnM*、*trnF*、*trnP*、*trnT*、*trnW*、*trnY* 和 *trnV*) 分别对应 18 种不同的氨基酸, 特殊的是丝氨酸 (Ser) 和亮氨酸 (Leu) 的 tRNA 分别有 2 个, 分别是 *trnL1* (CUN) 和 *trnL2* (UUR) 以及 *trnS1* (UCN) 和 *trnS2* (AGN)。tRNA 二级结构通常呈三叶草形, 含有一个 7 bp 的氨基酸接受臂, 一个 D 臂和 D 环, 一个反密码子臂和反密码子环, 一个 TΨC 臂和 TΨC

环,以及一个可变环。许多昆虫中 *trnS2* 的 D 臂缺失形成一个大的 D 环 (Sheffield *et al.*, 2008, Wei *et al.*, 2010c)。蕈蚋总科 Sciaroidea 中 tRNA 的 TΨC 颈环结构和氨基酸接受臂的 3' 区域全部缺失 (Beckenbach and Joy, 2009)。半闭弯尾姬蜂 *Diadegma semiclausum* 中部分 tRNA 具有较长的可变区 (Wei *et al.*, 2009)。rRNA 基因 *rrnL* 和 *rrnS* 分别编码大亚基 16S (16S rRNA) 和小亚基 12S (12S rRNA), 通过 G-C 和 A-U, 甚至不太稳定的 G-U 配对形成二级结构, 再和一些核糖体蛋白质结合到一起形成三级结构。

3 线粒体基因组的碱基组成

昆虫线粒体基因组中 4 种碱基 (A、T、C、G) 的组成比例并不均衡, 碱基 A 和 T 的含量高于 G 和 C, 而且大多数昆虫线粒体基因组上 A 的含量高于 T, C 的含量高于 G。碱基组成的差异通常用 A + T 含量、AT 偏斜和 GC 偏斜 4 个参数表示, 前者用于描述序列中 A 和 T 的比例, 后者分别描述 A 和 T 含量的差异以及 G 和 C 含量的差异。目前发现 A + T 含量最低的昆虫线粒体基因组来自蜉蝣 *Parafronurus youi*, A + T 的含量只有 57% (Zhang *et al.*, 2008b), 最高的来自半闭弯尾姬蜂 (膜翅目: 姬蜂科), 为 87.4% (Wei *et al.*, 2009)。全变态类昆虫 A + T 含量通常较高, 其中膜翅目高于其他目的种类 (Gotzek *et al.*, 2010)。膜翅目内部广腰亚目线粒体基因的 A + T 含量通常低于细腰亚目, 而广旗腹蜂 *Evania appendigaster* (细腰亚目: 旗腹蜂科) 线粒体基因组中 A + T 与广腰亚目的种类相似, 说明膜翅目中碱基组成变化规律在细腰亚目中更为复杂 (Wei *et al.*, 2010c)。

通过对 120 个昆虫的线粒体基因组进行对比和分析, 发现昆虫中膜翅目茧蜂科、虱目长角羽虱科和半翅目粉虱科 3 个类群的线粒体基因组碱基组成偏向性与大多数昆虫相比发生了颠倒, A 的含量低于 T, C 的含量低于 G。进一步结合基因组复制起点的二级结构、基因排列以及线粒体复制和转录方式的分析, 发现 G 与 C 之间的相对含量与复制起点的方向相关, 而 A 与 T 之间的相对含量因基因方向、复制起点方向以及密码子位置的变化而改变。碱基组成偏向性对研究线粒体基因组复制和转录的机理具有重要参考价值 (Wei *et al.*, 2010a)。

4 线粒体基因组的控制区

昆虫线粒体基因组的控制区通常位于 *rrnS* 和 *trnI* 之间的 N 链上。由于基因重排, 在一些种类中控制区的位置会发生变化。鳞翅目已经测定的种类的控制区均位于 *rrnS* 和 *trnQ* 之间, 但其在基因组中的相对位置没有发生改变 (Gong *et al.*, 2011)。蓟马线粒体基因组中有 2 个控制区, 分别位于 *cob* 和 *cox3* 内侧和 *atp6* 和 *nad5* 内侧 (Shao, 2003a)。膜翅目茧蜂科、虱目的长角羽虱科和半翅目的粉虱科 3 个类群的线粒体基因组中控制区均发生了倒置而位于 J 链上, 该现象在昆虫中至少独立进化了 3 次, 而且在类群内部具有保守性 (Wei *et al.*, 2010a)。

控制区通常是线粒体基因组中最大的非编码区。控制区的长度在不同物种以及不同个体之间均有差异。昆虫中报道的短的控制区只有 70 bp (Zhou *et al.*, 2007), 黑腹果蝇的为 4 601 bp (Lewis *et al.*, 1995), 在一些甲虫中控制区长达 9 ~ 13 kb (Zhang and Hewitt, 1997)。即使亲缘关系非常近的物种之间或同一物种不同品系之间, 控制区的差异可能会非常大。野桑蚕 *Bombyx mandarina* 青州品系的控制区长 495 bp, 而土浦 (Tsukuba) 品系的为 747 bp, 青州品系的长度与家蚕更相似 (Hu *et al.*, 2009)。

控制区的长度变异主要是重复单元的长度和数量变化引起的。控制区中富含 A 和 T, 与正常区域相比, 其可发生的碱基替换的类型有限, 而且序列中含有的少量 G 和 C 可能会因功能的限制进一步降低替换的速率, 因此控制区的碱基替换速率并不一定高于线粒体基因组中的其他区域 (Zhang and Hewitt, 1997), 这一结论已经在种内控制区与线粒体基因组的其他区域或单拷贝核基因的比较中得到验证 (Vila and Bjrkklund, 2004)。

控制区是线粒体基因组的复制和转录主要调控区 (Taanman, 1999)。由于控制区的变异较大, 很难通过序列同源性比对其进化与功能。通过比较分析, 在昆虫线粒体基因组的控制区中发现了 5 个与复制和转录有关的保守元素 (Zhang and Hewitt, 1997)。对直翅目中线粒体基因组的控制区进行比较分析发现, 控制区作为保守元素之一的颈环结构在类群内部具有一定的保守性 (Zhao *et al.*, 2010, 2011)。但是, 随着更多物种

控制区的测定与比较,发现昆虫中控制区的结构并不保守 (Duenas *et al.*, 2006),甚至在许多种类中都没有发现这些按一定顺序排列的保守元素 (Wei *et al.*, 2010c)。

5 昆虫线粒体基因的重排

线粒体基因组中基因的排列在类群中非常保守,亚库巴果蝇 *Drosophila yakuba* 线粒体基因的排列被认为是昆虫中最原始的排列方式 (图 1),大多数昆虫线粒体基因均按此方式排列。在进化过程中,某些类群中基因的位置和方向经历了重排事件而发生改变 (Boore, 1999; Dowton *et al.*, 2002)。

线粒体基因的重排类型有多种划分标准。根

据重排基因的类型分为大片段重排 (major rearrangement),即蛋白质编码基因或 rRNA 基因的重排,以及小片段重排 (minor rearrangement),即 tRNA 基因重排 (Cameron *et al.*, 2007a);根据重排后基因的位置可以分为 (1) 移位 (translocation),即基因跨越蛋白质或 rRNA 基因重排到另外一个位置,基因的方向不变; (2) 倒置或原位倒置 (local inversion),即基因的相对位置不变,但方向发生改变,在膜翅目中原位倒置可以占到重排事件的 1/3 (Dowton and Austin, 1999); (3) 基因洗牌 (gene shuffling),即某个区域的基因发生顺序错乱的重排; (4) 异位倒置 (remote inversion),即基因的方向和相对位置都发生改变 (Dowton *et al.*, 2003)。

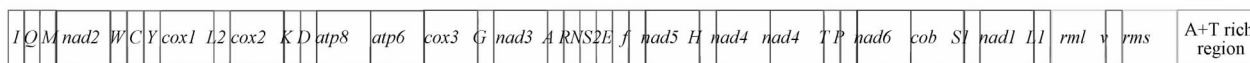


图 1 昆虫线粒体基因原始排列方式

Fig. 1 Putative ancestral mitochondrial gene arrangement of insects

方框底边加粗的基因编码于 N 链,未加粗的编码于 J 链。

Box with bold underscore indicates that the gene is encoded on minority strand while that with normal underscore indicates that the gene is encoded on majority strand.

有 14 个目的昆虫线粒体基因组发生重排。根据重排事件在类群内发生的频率及涉及到的基因数量,可以将昆虫中发生基因重排的目分为 3 类: (1) 基因高度重排的目,这些目中所有已经测定的种类均发生了基因重排事件,发生重排的基因数量多,除了 tRNA 基因外还可能涉及蛋白质编码基因和 rRNA 基因,如原尾目、虱目、缨翅目、啮虫目和膜翅目等 (Dowton and Austin, 1999, Shao *et al.*, 2001b; Dowton *et al.*, 2002, 2003); (2) 基因中度重排的目,这些目中部分种类发生基因重排,或全部种类发生少数几个基因的重排,如半翅目、双翅目、直翅目、鞘翅目、捻翅目、弹尾目、石蛎目; (3) 重排事件为共有衍征的目,这些目中所有种类均发生了相同的基因重排事件,重排事件可能是特定类群的共有衍征,如鳞翅目和脉翅目。基因高度重排的目主要集中在副新翅类 (半翅类) 中,而除了原尾目外另一高度重排的目—膜翅目—又位于副新翅类的姐妹群内翅类的基部位置。在六足总纲内,基因重排事件在各个目间是

随机发生的,与类群的进化关系没有直接的相关性 (Cameron *et al.*, 2006)。

5.1 基因高度重排的目

Chen 等 (2011) 测定了第 1 个原尾目昆虫线粒体基因组序列,发现 1 个蛋白质编码基因、2 个 rRNA 和 8 个 tRNA 基因发生重排。通过对另外 3 个原尾目种类进行检测,发现 *trnL2* 均由原始的位于 *cox1* 和 *cox2* 之间的位置重排到 *rnl* 和 *nad1* 之间。

虱目是基因重排最为频繁的目。*Heterodoxus macropus* 中所有的 tRNA 和 9 个蛋白质编码基因发生了重排 (Shao *et al.*, 2001a)。该目中编码于 N 链上的基因通常被重排到 J 链。*Campanulotes bidentatus compar* 经重排后,除 *trnQ* 外所有基因均编码于 J 链 (Covacin *et al.*, 2006); *Bothriometopus macrocnemis* 经重排后,所有基因均编码于 J 链,这是昆虫中发现的唯一一个所有线粒体基因编码于同一条链的种类 (Cameron *et al.*, 2007a)。虱目的同一科内基因排列顺序没有差异或差异较小,

不同亚目之间基因排列顺序差异巨大,几乎不存在相同排列顺序的区域 (Covacin *et al.*, 2006)。有些虱子的线粒体基因通过重排而形成一些独立的小环(minicircle),在已经报道的 10 个种类中 6 个种存在小环基因组。小环基因组可以分为 3 类,第一类是与完整的包含典型的线粒体基因的基因组共存的小环,小环上的基因与完整的基因组上的基因重复存在;第二类是包含多个基因和一个短的简单控制区的小环;第三类是包含一个或多个基因以及一个长的复杂控制区的小环。这些小环基因组的产生可能与核基因组中缺失编码线粒体单链结合蛋白(mitochondrial single-stranded binding protein)的基因有关 (Cameron *et al.*, 2011)。

缨翅目中测定的澳洲疫蓟马 *Thrips imaginis* 线粒体基因组中 24 个基因发生了重排,存在 2 个控制区,其中 2 个 rRNA 均发生了异位倒置且相距较远 (Shao, 2003a)。由于缨翅目中仅测定了一个种类的线粒体基因组序列,该目内部基因重排的频率及类群间的差异性尚未知。

啮虫目相对于虱目和缨翅目基因重排频率稍低,其 2 个亚目的种类之间基因排列顺序各不相同 (Shao, 2001b)。

膜翅目所有测定的 32 个线粒体基因组均发生了基因重排。基于线粒体基因组中的部分区域以及全基因组序列的比较均表明细腰亚目基因重排频率明显高于广腰亚目 (Dowton and Austin, 1999; Dowton *et al.*, 2003; Castro and Dowton, 2005)。膜翅目基因重排的类型非常丰富,移位、原位倒置、异位倒置和基因洗牌均有发生 (Dowton and Austin, 1999; Dowton *et al.*, 2003)。膜翅目中发生重排的主要是 tRNA 基因,控制区与 *nad2*、*nad2* 与 *cox1*、*cox2* 与 *atp8* 和 *nad3* 与 *nad5* 的基因链接部是 tRNA 重排的多发区。大多数 tRNA 基因重排事件都是独立发生的,其位置在进化过程中不受选择作用的影响。除了 tRNA 重排外,发现菜蛾盘绒茧蜂 *Cotesia vestalis* 和 3 种金小蜂属 *Nasonia* 的种类线粒体基因组中发生了蛋白质编码基因的重排 (Oliveira *et al.*, 2008; Wei *et al.*, 2010b),双色麦蜂 *Melifera bicolor* 中发生了 rRNA 的倒置 (Silvestre and Arias, 2006)。

5.2 基因中度重排的目

半翅目多数种类的基因都位于保守的位置,重排事件在目内不具有普遍性,但是多个总科和科内基因重排事件有加速的趋势。相对于半翅类另外 3 个目,半翅目重排的基因数量较少,以 tRNA 基因重排为主。胸喙亚目 Sternorrhyncha 中测定的 6 个粉虱、2 个蚜虫和 1 个木虱的线粒体基因组中,蚜虫、木虱和 2 种粉虱中发生了 1~2 个 tRNA 重排,另外 4 种粉虱中除了 tRNA 基因重排外,还有一段蛋白质编码基因区域(*cox3-trnG-nad3*)发生了异位倒置,重排到了 *rrnS* 和 *nad2* 之间。粉虱科内部系统发育分析表明,基因排列顺序相对保守的种类位于基部位置,而重排频率较高的种类位于较为进化的位置 (Thao *et al.*, 2004);古喙亚目 Archaeorrhyncha 4 个蜡蝉总科 Fulgoroidea 种类的线粒体基因组中,灰飞虱 *Laodelphax striatellus* 的 5 个 tRNA 和 3 个蛋白质编码基因发生了重排 (Song and Liang, 2009);前喙亚目 Prosorrhyncha 中测定的 40 个线粒体基因组中,扁蝽总科 Aradoidea 1 个种和红蝽总科 Pyrrhocoroidea 2 个种的线粒体基因组中均发生了 tRNA 重排现象,其中 *trnT* 和 *trnP* 的重排支持了大红蝽科 Largidae 和红蝽科 Pyrrhocoridae 的姐妹群关系 (Hua *et al.*, 2008, 2009)。

直翅目蝗亚目 Caelifera 中 Acridomorpha 所有种类中 *trnK-trnD* 发生重排形成 *trnD-trnK* 排列方式,成为该类群中的一个共有衍征 (Sheffield *et al.*, 2010; Zhao *et al.*, 2010)。另外,黄脸油葫芦 *Teleogryllus emma* 发生 tRNA 重排形成 *trnA-trnR-trnE-trnS-trnN-trnF* 排列方式。剑尾亚目所有种类和蝗亚目中蚤螳总科 Tridactyloidea 的 1 个种类没有发生基因重排。

双翅目中发生基因重排的种类全部在长角亚目 Nematocera 中。库蠓属 *Culicoides* 内部分种类发生 tRNA 基因重排,但重排与种类的系统发育位置没有相关性 (Matsumoto *et al.*, 2009);蚊科 Culicidae 所有已经测定的种类中 *trnA* 和 *trnR* 发生重排形成 *trnR-trnA* 排列方式 (Krzywinski *et al.*, 2006);虻蚋总科 Sciarioidea 所有种类均发生基因重排 (Beckenbach and Joy, 2009)。

鞘翅目中仅发现花蚤 *Mordella atrata* tRNA 重排形成 *trnA-trnN-trnR-trnS-trnE-trnF* 的排列方式 (Cameron *et al.*, 2009),象甲 *Naupactus xanthographus* tRNA 重排形成 *trnR-trnA-trnN-trnS-*

trnE-trnF 的排列方式 (Song *et al.*, 2010)。

捻翅目的 3 个种类在 *nad3* 和 *nad5* 之间的 tRNA 区发生了基因重排 (Carapelli *et al.*, 2006; McMahon *et al.*, 2009)。

石蛎目中测定的 4 个线粒体基因组中 *Pedetontus silvestrii* 发生了 *trnQ* 倒置 (Zhang *et al.*, 2008a)。

弹尾目中部分类群发生了 tRNA 基因重排。长角跳总科 Entomobryoidea 的 2 个种没有发生重排;棘跳科 Onychiuridae 2 个种类中 *trnQ* 和 *trnS1* 发生重排,同一总科的另外 4 个种没有发生重排;圆跳科 Sminthuridae 1 个种中 *trnF* 和 *trnD* 发生重排。

5.3 重排事件为共有衍征的目

脉翅目中所有测定的 5 个种类中 *trnW* 和 *trnC* 发生了重排形成了 *trnC-trnW-trnY* 排列方式 (Beckenbach and Stewart, 2008; Haruyama *et al.*, 2010; Negrisolo *et al.*, 2011)。

鳞翅目中所有测定的 40 个线粒体基因组均发生了 *trnM* 重排,形成了 *trnM-trnI-trnQ* 的排列方式,但是目前已经测定的鳞翅目线粒体基因组均属于双孔次目 Ditrysia,基部原始类群线粒体基因的排列顺序仍然未知 (Gong *et al.*, 2011),低等蝙蝠蛾总科中该区域的 tRNA 仍然保留原始排列方式 (曹永强, 2010),*trnM-trnI-trnQ* 是从哪一个类群开始进化出来的仍需进行对更多原始鳞翅目种类进行测序 (Yukuhiro *et al.*, 2002)。

5.4 基因重排的机理

了解基因发生重排的机理是正确利用基因重排事件的基础,是研究基因组进化的重要内容。目前主要有 4 种解释线粒体基因重排机理的模型与假说:(1)重复/随机丢失模型 (tandem duplication followed by random loss model, TDRL),即基因被复制后,其中一个拷贝被随机删除 (San Mauro *et al.*, 2006);(2)由 tRNA 基因不正当引发的复制 (illicit priming of replication by tRNA genes) (Cantatore *et al.*, 1987);(3)重组 (recombination) (Lunt *et al.*, 1998; Dowton and Campbell, 2001);(4)重复/非随机丢失 (duplication/nonrandom loss) (Lavrov *et al.*, 2002)。

一般认为基因洗牌是由基因的重复/随机丢失引起的,这种情况下通常可以在基因组中找到

被删除基因的残余序列或失去功能的重复基因的序列 (Macey *et al.*, 1998; Dowton and Austin, 1999)。膜翅目中发生重排的基因两侧通常具有较长的基因间隔区,而没有发生重排的基因两侧的间隔区通常较小或相邻基因之间存在重叠区 (Wei *et al.*, 2009),鳞翅目中发生重排的 *trnM* 的区域同样存在较大的非编码区 (Gong *et al.*, 2011)。造成基因移位的原因可能有多种,一种解释是重复/随机丢失,这种情况下需要有较长序列发生重复,然后大量重复的基因同时被删除。然而,发生大片段基因重复丢失的概率非常小。长翅目中发现了大约 4 kb 的包含蛋白质编码基因区域的重复序列,表明昆虫中发生大片段基因序列重复是可能的 (Beckenbach and Civetta, 2011)。基因组内部重组和 tRNA 引起的复制时的错误引导也会造成移位 (Boore and Brown, 1998; Dowton and Austin, 1999)。蓟马、虱目和膜翅目中发生的大量基因倒置,目前最合理的解释是基因组内重组 (intramitochondrial recombination) (Dowton and Campbell, 2001)。异位倒置是最难解释的重排事件,通过对膜翅目茧蜂科的 *trnH* 进行分析,推测异位倒置可能是经过 2 次独立的重组事件产生的 (Wei *et al.*, 2010b)。

5.5 基因重排加速的原因

为什么有些类群中会发生高频率的基因重排事件呢?

半翅类昆虫中,基因重排速率越高其碱基替换速率越高,二者的相关性具有统计显著性。在其他类群的昆虫中,基因重排较多的种类中亦存在相关性,而重排较少的种类中二者不存在相关性 (Shao, 2003b)。对节肢动物的分析表明,多数情况下基因的重排速率越高进化速率越高 (Xu *et al.*, 2006)。而对茧蜂科昆虫进行分析发现,重排严重的种类具有较高的进化速率,但是具有较高进化速率的种类基因重排速率不一定高 (Wei *et al.*, 2010b)。同样,半翅目的蝽次目 Pentatomomorpha 中,基因重排速率与碱基替换速率之间不具有相关性 (Hua *et al.*, 2008)。

在已经测通的线粒体基因组的数量较少时,初步认为基因重排与寄生等生物学习性有关。例如发生基因重排的虱子、蚊子和膜翅目昆虫等均具有寄生习性 (Dowton and Austin, 1999; Shao

et al., 2001a)。随着更多发生基因重排的线粒体基因组被测定,发现植食性的蓟马、粉虱等也具有较高的基因重排速率 (Shao, 2003a; Thao *et al.*, 2004)。膜翅目中,虽然以寄生为主的细腰亚目中基因重排的频率高于以植食性为主的广腰亚目,但是细腰亚目中也存在基因重排频率较低的种类,因此,膜翅目中基因重排速率与寄生习性之间不存在必然的联系 (Dowton *et al.*, 2009b)。

6 线粒体基因的进化速率

在过去 30 多年里线粒体基因之所以成为进化与多样性研究中最常用的分子标记,主要是因为其较快的进化速率、母性遗传、中性进化以及符合分子钟假说等特性 (Brown *et al.*, 1979; Curole and Kocher, 1999)。然而,许多研究对线粒体基因这些特性提出了质疑 (Ballard and Whitlock, 2004; Hurst and Jiggins, 2005; Galtier *et al.*, 2009)。随着大量线粒体基因组序列被测定以及基于全基因组层面的比较分析,我们对线粒体基因进化速率有了更加全面的认识。

通常线粒体基因的进化速率快于核基因。最特殊的例子来自 3 种金小蜂属 *Nasonia* 的种类,其线粒体基因组的进化速率是核基因的 30 倍 (Oliveira *et al.*, 2008);不同类群间线粒体基因进化速率存在差异。膜翅目广腰亚目的进化速率与其他全变态类昆虫相似,而细腰亚目的进化速率是广腰亚目的 2~3 倍 (Dowton *et al.*, 2009a)。对膜翅目茧蜂科 7 个亚科进化速率的比较分析发现,非圆口类的进化速率快于圆口类 (Wei *et al.*, 2010b);不同线粒体基因以及蛋白质编码基因的不同编码位点的进化速率存在差异。由于受二级结构的限制,tRNA 进化速率最慢,其次是 rRNA 基因,最快的是蛋白质编码基因。在 13 个蛋白质编码基因中,细胞色素 C 氧化酶亚基基因进化慢,而 NADH 脱氢酶亚基基因进化快。在蛋白质编码基因的 3 个编码位点中,第 3 位编码位点进化速率快于第 1 和第 2 位 (Cameron and Whiting, 2008; Hua *et al.*, 2008; Negrisolo *et al.*, 2011)。

7 线粒体基因组在昆虫系统发育研究中的应用

7.1 线粒体基因组中的分子标记

线粒体基因组中包含多种可用于系统发育分

析的分子标记。核酸序列和氨基酸序列是最为广泛应用的标记类型。控制区结构、基因排列顺序中可能含有对系统发育关系分析有用的信息 (Dowton *et al.*, 2002; Zhao *et al.*, 2010, 2011)。对直翅目控制区进行分析发现,其颈环结构可以很好的解决科内亚科间的系统发育关系,但无法解决科及更高阶分类阶元间的系统发育关系 (Zhao *et al.*, 2011)。虽然在昆虫中发现了大量基因重排事件,目前已经非常明确,重排事件在解决昆虫目之间的系统发育关系方面不能提供有用的信息 (Cameron *et al.*, 2006)。但是昆虫中频繁的基因重排事件可能为目内类群间系统发育关系的研究提供重要信息 (Cameron *et al.*, 2006)。rRNA 二级结构可以指导序列联配 (Cannone *et al.*, 2002; Gillespie *et al.*, 2006),同时也含有直接用于推测类群间的系统发育关系的信息 (Wei *et al.*, 2010b)。例如,在膜翅目茧蜂科 7 个亚科中,*rrnS* 颈环结构的变化规律与基于线粒体基因组序列构建的系统发育关系完全吻合 (Wei *et al.*, 2010b)。

7.2 线粒体基因组数据的系统发育分析方法

在应用线粒体基因组构建系统发育关系时,数据矩阵的处理以及系统发育运算方法至关重要。例如,研究膜翅目在全变态类昆虫中的位置时,随着选用的数据矩阵和系统发育运算方法的改变而得出不同的结论 (Castro and Dowton, 2007)。

由于蛋白质编码基因第 3 位密码子容易发生替换饱和,数据分析时常去除或用嘌呤和嘧啶重新编码该密码子位的碱基。对全变态类昆虫多个目进行系统发育分析时,两种处理方法均可以改善目之间以及目内主要类群间的系统发育关系的分析结果 (Wei *et al.*, 2010b)。对膜翅目进行分析时,去除蛋白质编码基因第 3 密码子位的碱基得到的系统发育结果与目前公认的系统发育关系更加一致 (Castro and Dowton, 2007; Dowton *et al.*, 2009a),对直翅目和半翅目进行分析时,保留第 3 密码子位的碱基没有降低系统发育结果的可靠性 (Fenn *et al.*, 2008; Hua *et al.*, 2008, 2009),而对双翅目内主要类群的系统发育关系的研究表明,去除蛋白质编码基因第 3 密码子位的碱基降低了系统发育分析结果的可靠性

(Cameron *et al.*, 2007b)。

另外一种数据矩阵处理方法是將蛋白质编码基因翻译成氨基酸序列。对全变态类昆虫进行分析时,基于蛋白质编码基因的氨基酸序列构建的贝叶斯系统发育树的拓扑结构与基于核酸序列得到的结果基本一致(Wei *et al.*, 2010b),而直翅目中基于氨基酸序列得到的拓扑结构中存在很多错误的系统发育关系(Fenn *et al.*, 2008)。由于基于氨基酸序列的分析运算量较大,因此,多数研究仅选用线粒体基因组中的核酸序列进行分析。

有些研究在构建数据矩阵时对不同基因进行取舍。对直翅目的分析表明,蛋白质编码基因对系统发育分析结果起到主要作用,能较好的解决高级阶元间的系统发育关系。rRNA 和 tRNA 单独用于系统发育分析时得到的结果非常不理想,受系统发育分析过程中所用参数的影响较大,而与蛋白质编码基因同时使用可以增强系统发育分析信号的强度(Fenn *et al.*, 2008; Zhao *et al.*, 2010)。双翅目中同样发现在蛋白质编码基因的基础上增加 rRNA 和 tRNA 可以提高数据矩阵解决系统发育关系的能力并提高分支的支持率(或后验概率)(Cameron *et al.*, 2007b)。半翅目中,单独使用蛋白质编码基因与增加 rRNA 和 tRNA 基因的矩阵得到相同的系统发育拓扑结构(Hua *et al.*, 2009)。对节肢动物线粒体基因进行分析发现,13 个蛋白质编码基因对系统发育分析结果的贡献是均等的,去除基因会明显的降低数据矩阵解决系统发育关系的能力(Cameron *et al.*, 2004)。脉翅类线粒体基因组中 *cob*、*cox1* 和 *cox3* 的系统发育信号最好,其次是 *atp6*、*cox2*、*nad4*、*nad5* 和 *nad1*,最差的是 *nad2*、*nad3* 和 *nad6* (Negrisolo *et al.*, 2011)。目前,多数研究通常选择 13 个蛋白质编码基因的第 1 和 2 为编码位点的碱基进行系统发育分析。在利用线粒体基因组构建系统发育关系时,建议包含蛋白质编码基因、tRNA 和 rRNA 基因(Cameron *et al.*, 2006)。

对线粒体基因组的数据矩阵进行分区(partition)通常可以优化位点之间的似然值。实际操作时可以根据不同的基因或蛋白质编码基因的不同编码位点进行分区。在直翅目、双翅目、膜翅目等类群的研究中均证实了这一点(Cameron *et al.*, 2007b; Fenn *et al.*, 2008; Downton *et al.*, 2009a)。但是过度分区会起到相反的作用。当对

数据矩阵同时按照基因和编码位点进行分区时,系统发育运算结果很难趋同。这可能是因为过多的分区使得运算过程中难以对多个小的分区的参数进行准确的估算(Fenn *et al.*, 2008)。

外群的选择也会对系统发育结果造成影响(Cameron *et al.*, 2004; Castro and Downton 2007)。多数研究选择 2 个以上的外群以提高系统发育分析结果的稳定性(Cameron *et al.*, 2007b; Downton *et al.*, 2009a; Wei *et al.*, 2010b)。

贝叶斯法、最大似然法和最大简约法是构建系统发育关系时最常用的运算方法。虽然在一些类群如双翅目内部,3 种运算方法得到的拓扑结构完全一致(Cameron *et al.*, 2007b),但是多数研究表明贝叶斯法得到的结果优于最大似然法和最大简约法(Downton *et al.*, 2009a; Wei *et al.*, 2010b)。贝叶斯法虽然可以在一定程度上克服碱基组成异质性的影响,但是结果仍然会受到位点间变异速率异质性的影响(Song *et al.*, 2010)。

7.3 应用线粒体基因组构建的昆虫系统发育关系

线粒体基因组被广泛用于昆虫在节肢动物中的位置、昆虫的目、总科、科以及亚科等不同分类阶元层次的系统发育关系研究。

在昆虫的单系性及其在节肢动物中的位置方面,Boore 等(1998)根据一个 tRNA 的重排提出昆虫与甲壳纲为姐妹群关系,这一观点得到了分子和形态学数据的支持(Giribet *et al.*, 2001)。基于线粒体基因组基因序列构建的系统发育关系普遍认为昆虫是一个并系群(Nardi *et al.*, 2001, 2003, Cook *et al.*, 2005; Carapelli *et al.*, 2007; Chen *et al.*, 2011)。Nardi 等(2003)利用 35 个昆虫线粒体基因组分析时发现弹尾目位于甲壳纲和昆虫之外。Chen 等(2011)则发现弹尾目位于昆虫内部,而原尾目位于昆虫外部,根据 *cox1-trnL2-cox2* 排列方式在泛甲壳类 Pancrustacean(甲壳类 Crustacea 和六足总纲 Hexapoda)中的保守性,以及原尾目 *cox1-cox2* 排列方式,推测六足总纲的单系性不成立,而原尾目可能是早于其他泛甲壳类而进化出来的一个分支(Chen *et al.*, 2011)。由于线粒体基因进化方式的复杂性,选用的数据处理和分析方法的局限性以及构建系统发育树所用的代表性类群的局限性等原因,基于线粒体基因组

序列判断昆虫是否是一个单系群尚需进一步研究 (Delsuc *et al.*, 2003; Cook *et al.*, 2005)。

在目以上分类单元间的系统发育关系研究中,支持了蜚蠊目(包含等翅目)和螳螂目的姐妹群关系 (Zhang *et al.*, 2009)。全变态类昆虫(内翅部)作为昆虫中最为高等、种类最为繁多的类群,其内部各目之间的系统发育关系争议较大。基于线粒体基因组的系统发育研究支持膜翅目位于全变态类昆虫的基部位置 (Wei *et al.*, 2010b),与利用其他类型的分子标记得出的结论相一致 (Savard *et al.*, 2006; Misof *et al.*, 2007; Wiegmann *et al.*, 2009)。鳞翅目、鞘翅目、蛇蛉目、脉翅目、广翅目、长翅目和双翅目的位置仍不稳定 (Cameron *et al.*, 2009; Wei *et al.*, 2010b),与利用其他类型的分子标记得出的结论之间仍然存在差异 (Wiegmann *et al.*, 2009)。

目前已经基于线粒体基因组构建了昆虫多个目内主要类群间的系统发育关系。分析结果与基于其他标记构建的系统发育关系较为一致的目有直翅目 (Fenn *et al.*, 2008; Zhou *et al.*, 2010)、双翅目 (Cameron *et al.*, 2007b)、半翅目 (Hua *et al.*, 2008, 2009)和鳞翅目 (Kim *et al.*, 2011)等。对膜翅目的分析支持广腰亚目是一个并系类群以及细腰亚目的单系性,但是细腰亚目内部类群间的系统发育关系仍不稳定 (Dowton *et al.*, 2009a; Gotzek *et al.*, 2010; Wei *et al.*, 2010b),这可能与膜翅目内各类群线粒体基因组的进化速率差异较大有关 (Dowton *et al.*, 2009a)。鞘翅目线粒体基因组存在碱基组成异质性以及位点间分化速率异质性,因此构建的系统发育关系受数据处理和运算方法影响较大 (Song *et al.*, 2010)。

8 结论与展望

目前测定的线粒体基因组序列涵盖了昆虫中大多数目,以及直翅目、半翅目、膜翅目、鳞翅目、鞘翅目、双翅目等几个种类繁多的目内部的主要类群。这些序列为比较基因组学提供了丰富的数据资料。然而,相对于昆虫这一庞大的类群,这些测定的线粒体基因组序列仍然非常有限。例如,鳞翅目中测定的线粒体基因组均属于双孔次目 Dityisia,基部原始类群极少涉及;缨翅目中报道的一个种类的线粒体基因组发生了频繁的基因重排事件,而基因重排事件在目内的规律仍然未知。

因此,今后仍需对一些重要类群和重要种类的线粒体基因组进行测序和分析。

过去测定的线粒体基因组较少时,研究者仅对新测定的基因组进行结构和特征描述。近些年,针对某一类特定类群或某一现象的比较逐渐增多,从而对线粒体基因组的进化趋势和特征有了更加深入的了解。今后,仍需针对基因重排规律、基因重排产生的原因、碱基组成、进化速率等方面的问题,选择一些合适的研究类群,进行系统的比较分析。

目前的研究已经表明线粒体基因组在构建昆虫系统发育关系方面具有巨大的潜力。而昆虫线粒体基因组特殊的进化方式,如碱基组成异质性、类群间、基因间以及位点间进化速率的异质性、部分类群内频繁的基因重排事件等,一方面给线粒体基因组在系统发育研究中的应用带来了机遇,为开发更多有用的信息提供了数据来源,另一方面为正确利用线粒体基因组带来了挑战。因此,在深入了解线粒体基因组进化方式与特征的基础上,选择合适的系统发育分析方法是构建可靠的系统发育关系的关键。

参考文献 (References)

- Ballard JWO, Whitlock MC, 2004. The incomplete natural history of mitochondria. *Mol. Ecol.*, 13(4):729—744.
- Beckenbach AT, Civetta A, 2011. Mitochondrial genome sequences of representatives of three families of scorpionflies (Order Mecoptera) and evolution in a major duplication of coding sequence. *Genome*, 54(5):368—376.
- Beckenbach AT, Joy JB, 2009. Evolution of the mitochondrial genomes of gall midges (Diptera: Cecidomyiidae): rearrangement and severe truncation of tRNA genes. *Genome Biol. Evol.*, 2009(1):278—287.
- Beckenbach ATBAT, Stewart JBSJB, 2008. Insect mitochondrial genomics 3: the complete mitochondrial genome sequences of representatives from two neuropteroid orders: a dobsonfly (order Megaloptera) and a giant lacewing and an owlfly (order Neuroptera). *Genome*, 52(1):31—38.
- Boore JL, 1999. Animal mitochondrial genomes. *Nucleic Acids Res.*, 27(8):1767—1780.
- Boore JL, Brown WM, 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. *Curr. Opin. Genet. Dev.*, 8(6):668—674.

- Boore JL, Lavrov DV, Brown WM, 1998. Gene translocation links insects and crustaceans. *Nature*, 392(6677):667—668.
- Brown WM, George M, Wilson AC, 1979. Rapid evolution of animal mitochondrial DNA. *PNAS*, 76(4):1967—1971.
- Cameron S, Yoshizawa K, Mizukoshi A, Whiting M, Johnson K, 2011. Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). *BMC Genomics*, 12(1):394.
- Cameron SL, Beckenbach AT, Dowton M, Whiting M, 2006. Evidence from mitochondrial genomics on interordinal relationships in insects. *Arthropod Syst. Phylog.*, 64(1):27—34.
- Cameron SL, Dowton M, Castro LR, Ruberu K, Whiting MF, Austin AD, Diement K, Stevens J, 2008. Mitochondrial genome organization and phylogeny of two vespid wasps. *Genome*, 51(10):800—808.
- Cameron SL, Johnson KP, Whiting MF, 2007a. The mitochondrial genome of the screamer louse *Bothriometopus* (Phthiraptera: Ischnocera): effects of extensive gene rearrangements on the evolution of the genome. *J. Mol. Evol.*, 65(5):589—604.
- Cameron SL, Lambkin CL, Barker SC, Whiting MF, 2007b. A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. *Syst. Entomol.*, 32(1):40—59.
- Cameron SL, Miller KB, D'Haese CA, Whiting MF, Barker SC, 2004. Mitochondrial genome data alone are not enough to unambiguously resolve the relationships of Entognatha, Insecta and Crustacea sensu lato (Arthropoda). *Cladistics*, 20(6):534—557.
- Cameron SL, Sullivan J, Song H, Miller KB, Whiting MF, 2009. A mitochondrial genome phylogeny of the Neuropterida (lace-wings, alderflies and snakeflies) and their relationship to the other holometabolous insect orders. *Zool. Scr.*, 38(6):575—590.
- Cameron SL, Whiting MF, 2008. The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene*, 408(1/2):112—123.
- Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Muller KM, 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, 3(2):1471—2105.
- Cantatore P, Gadaleta MN, Roberti M, Saccone C, Wilson AC, 1987. Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. *Nature*, 329(6142):853—855.
- 曹永强, 2010. 人支蝠蛾 (*Hepialus renzhiensis*) 和云南蝠蛾 (*Hepialu synnanensis*) 的线粒体基因组结构及蝠蛾属昆虫分子系统发育研究. 博士学位毕业论文. 西双版纳: 中国科学院西双版纳热带植物园.
- Carapelli A, Liò P, Nardi F, Van Der Wath E, Frati F, 2007. Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evol. Biol.*, 7(Suppl 2):S8.
- Carapelli A, Vannini L, Nardi F, Boore JL, Beani L, Dallai R, Frati F, 2006. The mitochondrial genome of the entomophagous endoparasite *Xenos vesparum* (Insecta: Strepsiptera). *Gene*, 376(2):248—259.
- Castro L, Dowton M, 2005. The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera: Symphyta: Pergidae). *Mol. Phylogenet. Evol.*, 34(3):469—479.
- Castro LR, Dowton M, 2007. Mitochondrial genomes in the Hymenoptera and their utility as phylogenetic markers. *Syst. Entomol.*, 32(1):60—69.
- Castro LR, Ruberu K, Dowton M, 2006. Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: Vanhorniidae) and *Primeuchroeus* spp. (Aculeata: Chrysididae): evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera). *Genome*, 49(7):752—766.
- Chen WJ, Bu Y, Carapelli A, Dallai R, Li S, Yin WY, Luan YX, 2011. The mitochondrial genome of *Sinentomon erythranum* (Arthropoda: Hexapoda: Protura): an example of highly divergent evolution. *BMC Evol. Biol.*, 11(1):246.
- Cook CE, Yue Q, Akam M, 2005. Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proc. R. Soc. Biol. Sci. Ser. B*, 272(1569):1295.
- Covacin C, Shao R, Cameron S, Barker SC, 2006. Extraordinary number of gene rearrangements in the mitochondrial genomes of lice (Phthiraptera: Insecta). *Insect Mol. Biol.*, 15(1):63—68.
- Curole JP, Kocher TD, 1999. Mitogenomics: digging deeper with complete mitochondrial genomes. *Trends Ecol. Evol.*, 14(10):394—398.
- Delsuc F, Phillips MJ, Penny D, 2003. Comment on "Hexapod origins: monophyletic or paraphyletic?". *Science*,

- 301(5639):1482.
- Dowton M, Austin AD, 1999. Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. *Mol. Biol. Evol.*, 16(2):298—309.
- Dowton M, Cameron SL, Austin AD, Whiting MF, 2009a. Phylogenetic approaches for the analysis of mitochondrial genome sequence data in the Hymenoptera-A lineage with both rapidly and slowly evolving mitochondrial genomes. *Mol. Phylogenet. Evol.*, 52(2):512—519.
- Dowton M, Cameron SL, Dowavic JI, Austin AD, Whiting MF, 2009b. Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. *Mol. Biol. Evol.*, 26(7):1607—1617.
- Dowton M, Campbell NJH, 2001. Intramitochondrial recombination-is it why some mitochondrial genes sleep around? *Trends. Ecol. Evol.*, 16(6):269—271.
- Dowton M, Castro LR, Austin AD, 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. *Invertebr. Syst.*, 16(1):345—356.
- Dowton M, Castro LR, Campbell SL, Bargon SD, Austin AD, 2003. Frequent mitochondrial gene rearrangements at the Hymenopteran nad3 – nad5 junction. *J. Mol. Evol.*, 56(5):517—526.
- Duenas JCR, Gardenal CN, Llinas GA, Panzetta-Dutari GM, 2006. Structural organization of the mitochondrial DNA control region in *Aedes aegypti*. *Genome*, 49(8):931—937.
- Fenn JD, Song H, Cameron SL, Whiting MF, 2008. A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. *Mol. Phylogenet. Evol.*, 49(1):59—68.
- Galtier N, Nabholz B, Glémin S, Hurst G, 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol. Ecol.*, 18(22):4541—4550.
- Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR, 2006. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Mol. Biol.*, 15(5):657—686.
- Gribet G, Edgecombe GD, Wheeler WC, 2001. Arthropod phylogeny based on eight molecular loci and morphology. *Nature*, 413(6852):157—161.
- Gong YJ, Shi BC, Kang Z, J, , Zhang F, Wei SJ, 2011. The complete mitochondrial genome of the oriental fruit moth *Grapholita molesta* (Busck) (Lepidoptera: Tortricidae). *Mol. Biol. Rep.*:DOI 10.1007/s11033—11011—11049—y.
- Gotzek D, Clarke J, Shoemaker D, 2010. Mitochondrial genome evolution in fire ants (Hymenoptera: Formicidae). *BMC Evol. Biol.*, 10(1):300.
- Haruyama N, Mochizuki A, Sato Y, Naka H, Nomura M, 2010. Complete mitochondrial genomes of two green lacewings, *Chrysoperla nipponensis* (Okamoto, 1914) and *Apochrysa matsumurae* Okamoto, 1912 (Neuroptera: Chrysopidae). *Mol. Biol. Rep.*, 38(5):57—61.
- Hu XL, Cao GL, Xue RY, Zheng XJ, Zhang X, Duan HR, Gong CL, 2009. The complete mitogenome and phylogenetic analysis of *Bombyx mandarina* strain Qingzhou. *Mol. Biol. Rep.*, 37(6):2599—2608.
- Hua JM, Li M, Dong PZ, Cui Y, Xie Q, Bu WJ, 2008. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera). *BMC Genomics*, 9(1):610.
- Hua JM, Li M, Dong PZ, Cui Y, Xie Q, Bu WJ, 2009. Phylogenetic analysis of the true water bugs (Insecta: Hemiptera: Heteroptera: Nepomorpha): evidence from mitochondrial genomes. *BMC Evol. Biol.*, 9(1):134.
- Hurst GDD, Jiggins FM, 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. R. Soc. Biol. Sci. Ser. B*, 272(1572):1525—1534.
- Kim MJ, Kang AR, Jeong HC, Kim KG, Kim I, 2011. Reconstructing intraordinal relationships in Lepidoptera using mitochondrial genome data with the description of two newly sequenced lycaenids, *Spindasis takanonis* and *Protantigius superans* (Lepidoptera: Lycaenidae). *Mol. Phylogenet. Evol.*, 2(61):436—445.
- Krzywinski J, Grushko OG, Besansky NJ, 2006. Analysis of the complete mitochondrial DNA from *Anopheles funestus*: An improved dipteran mitochondrial genome annotation and a temporal dimension of mosquito evolution. *Mol. Phylogenet. Evol.*, 39(2):417—423.
- Lavrov DV, Boore JL, Brown WM, 2002. Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. *Mol. Biol. Evol.*, 19(2):163—169.
- Lewis DL, Farr CL, Kaguni LS, 1995. *Drosophila melanogaster* mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons. *Insect*

- Mol. Biol.*, 4(4):263—278.
- Lunt DH, Whipple LE, Hyman BC, 1998. Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. *Mol. Ecol.*, 7(11):1441—1455.
- Macey JR, Schulte JA, Larson A, Papenfuss TJ, 1998. Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement. *Mol. Biol. Evol.*, 15(1):71—75.
- Matsumoto Y, Yanase T, Tsuda T, Noda H, 2009. Species-specific mitochondrial gene rearrangements in biting midges and vector species identification. *Med. Vet. Entomol.*, 23(1):47—55.
- McMahon DP, Hayward A, Kathirithamby J, 2009. The mitochondrial genome of the 'twisted-wing parasite' *Mengenilla australiensis* (Insecta, Strepsiptera): a comparative study. *BMC Genomics*, 10(1):603.
- Misof B, Niehuis O, Bischoff I, Rickert A, Erpenbeck D, Staniczek A, 2007. Towards an 18S phylogeny of hexapods: accounting for group-specific character covariance in optimized mixed nucleotide/doublet models. *Zoology*, 110(5):409—429.
- Nardi F, Carapelli A, Fanciulli PP, Dallai R, Frati F, 2001. The complete mitochondrial DNA sequence of the basal hexapod *Tetradontophora bielensis*: evidence for heteroplasmy and tRNA translocations. *Mol. Biol. Evol.*, 18(7):1293.
- Nardi F, Spinsanti G, Boore JL, Carapelli A, Dallai R, Frati F, 2003. Hexapod origins: Monophyletic or paraphyletic? *Science*, 299(5614):1887—1889.
- Nass MMK, Nass S, 1963. Intramitochondrial fibers with DNA characteristics. *J. Cell Biol.*, 19(3):593—611.
- Negrisola E, Babbucci M, Patarnello T, 2011. The mitochondrial genome of the ascalaphid owlfly *Libelloides macaronius* and comparative evolutionary mitochondriomics of neuropterid insects. *BMC Genomics*, 12(1):221.
- Oliveira DCSG, Raychoudhury R, Lavrov DV, Werren JH, 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Mol. Biol. Evol.*, 25(10):2167—2180.
- San Mauro D, Gower DJ, Zardoya R, Wilkinson M, 2006. A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. *Mol. Biol. Evol.*, 23(1):227—234.
- Savard J, Tautz D, Richards S, Weinstock GM, Gibbs RA, Werren JH, Tettelin H, Lercher MJ, 2006. Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of holometabolous insects. *Genome Res.*, 16(11):1334.
- Shao R, 2003a. The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. *Mol. Biol. Evol.*, 20(3):362—370.
- Shao R, 2003b. Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. *Mol. Biol. Evol.*, 20(10):1612—1619.
- Shao R, Campbell NJH, Barker SC, 2001a. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). *Mol. Biol. Evol.*, 18(5):858—865.
- Shao R, Kirkness EF, Barker SC, 2009. The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. *Genome Res.*, 19(5):904.
- Shao RF, Campbell NJH, Schmidt ER, Barker SC, 2001b. Increased rate of gene rearrangement in the mitochondrial genomes of three orders of hemipteroid insects. *Mol. Biol. Evol.*, 18(9):1828—1832.
- Sheffield NC, Hiatt KD, Valentine MC, Song H, Whiting MF, 2010. Mitochondrial genomics in Orthoptera using MOSAS. *Mitochondr. DNA*, 21(3/4):87—104.
- Sheffield NC, Song H, Cameron SL, Whiting MF, 2008. A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Mol. Biol. Evol.*, 25(11):2499—2509.
- Silvestre D, Arias MC, 2006. Mitochondrial tRNA gene translocations in highly eusocial bees. *Genet. Mol. Biol.*, 29(3):572—575.
- Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT, 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu. Rev. Ecol. Syst.*, 37(1):545—579.
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P, 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.*, 87(6):651—701.
- Song H, Sheffield NC, Cameron SL, Miller KB, Whiting MF, 2010. When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics. *Syst. Entomol.*, 35

- (1):429—448.
- Song N, Liang AP, 2009. Complete mitochondrial genome of the small brown planthopper, *Laodelphax striatellus* (Delphacidae:Hemiptera), with a novel gene order. *Zool. Sci.*, 26(12):851—860.
- Taanman JW, 1999. The mitochondrial genome: structure, transcription, translation and replication. *Biochim. Biophys. Acta*, 1410(2):103—123.
- Thao ML, Baumann L, Baumann P, 2004. Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). *BMC Evol. Biol.*, 4(1):25.
- Vila M, Bj rklund M, 2004. The utility of the neglected mitochondrial control region for evolutionary studies in Lepidoptera (Insecta). *J. Mol. Evol.*, 58(3):280—290.
- Wei SJ, Shi M, Chen XX, Sharkey MJ, van Achterberg C, Ye GY, He JH, 2010a. New views on strand asymmetry in insect mitochondrial genomes. *PLoS ONE*, 5(9):1767—1780.
- Wei SJ, Shi M, He JH, Sharkey MJ, Chen XX, 2009. The complete mitochondrial genome of *Diadegma semiclausum* (Hymenoptera: Ichneumonidae) indicates extensive independent evolutionary events. *Genome*, 52(1):308—319.
- Wei SJ, Shi M, Sharkey MJ, van Achterberg C, Chen XX, 2010b. Comparative mitogenomics of *Braconidae* (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to Holometabolous insects. *BMC Genomics*, 11(1):371.
- Wei SJ, Tang P, Zheng LH, Shi M, Chen XX, 2010c. The complete mitochondrial genome of *Evania appendigaster* (Hymenoptera: Eviiniidae) has low A + T content and a long intergenic spacer between atp8 and atp6. *Mol. Biol. Rep.*, 37(4):1931—1942.
- Wiegmann B, Trautwein M, Kim JW, Cassel B, Bertone M, Winterton S, Yeates D, 2009. Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biol.*, 7(1):34.
- Wilson AC, Cann RL, Carr SM, George M, Gyllenstein U, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.*, 26(4):375—400.
- Wolstenholme DR. 1992. Animal mitochondrial DNA: structure and evolution. *Int. Rev. Cytol.*, 141(1):173—216.
- Xu W, Jameson D, Tang B, Higgs PG, 2006. The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. *J. Mol. Evol.*, 63(3):375—392.
- Yukuhiro K, Sezutsu H, Itoh M, Shimizu K, Banno Y, 2002. Significant levels of sequence divergence and gene rearrangements have occurred between the mitochondrial genomes of the wild mulberry silkworm, *Bombyx mandarina*, and its close relative, the domesticated silkworm, *Bombyx mori*. *Mol. Biol. Evol.*, 19(8):1385—1389.
- Zhang DX, Hewitt GM, 1997. Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochem. Syst. Ecol.*, 25(2):99—120.
- Zhang JY, Song DX, Zhou KY, 2008a. The complete mitochondrial genome of the bristletail *Pedotontus silvestrii* (Archaeognatha: Machilidae) and an examination of mitochondrial gene variability within four bristletails. *Ann. Entomol. Soc. Am.*, 101(6):1131—1136.
- Zhang JY, Zhou CF, Gai YH, Song DX, Zhou KY, 2008b. The complete mitochondrial genome of *Parafonurus youi* (Insecta: Ephemeroptera) and phylogenetic position of the Ephemeroptera. *Gene*, 424(1/2):18—24.
- Zhang YY, Xuan WJ, Zhao JL, Zhu CD, Jiang GF, 2009. The complete mitochondrial genome of the cockroach *Eupolyphaga sinensis* (Blattaria: Polyphagidae) and the phylogenetic relationships within the Dictyoptera. *Mol. Biol. Rep.*, 37(1):3509—3516.
- Zhao L, Zheng ZM, Huang Y, Sun HM, 2010. A comparative analysis of mitochondrial genomes in *Orthoptera* (Arthropoda: Insecta) and genome descriptions of three grasshopper species. *Zool. Sci.*, 27(8):662—672.
- Zhao L, Zheng ZM, Huang Y, Zhou ZJ, Wang L, 2011. Comparative analysis of the mitochondrial control region in *Orthoptera*. *Zool. Stud.*, 50(3):385—393.
- Zhou ZJ, Huang Y, Shi FM, 2007. The mitochondrial genome of *Ruspolia dubia* (Orthoptera: Conocephalidae) contains a short A + T - rich region of 70 bp in length. *Genome*, 50(9):855—866.
- Zhou ZJ, Ye HY, Huang Y, Shi FM, 2010. The phylogeny of *Orthoptera* inferred from mtDNA and description of *Elimaeca cheni* (Tettigoniidae: Phaneropterinae) mitogenome. *J. Genet. Genomics*, 37(5):315—324.

昆虫生物钟分子调控研究进展^{*}

陈文锋 贺春霞 安春菊 赵章武^{**}

(中国农业大学农业与生物技术学院昆虫系 北京 100193)

摘 要 昆虫生物钟节律的研究是人类了解生物节律的重要途径。昆虫在生理和行为上具有广泛的节律活动,如运动、睡眠、学习记忆、交配、嗅觉等节律活动,其中昼夜活动行为为节律的研究广泛而深入。昆虫乃至高等动物普遍具有保守的昼夜节律系统,昼夜生物钟节律主要包括输入系统:用于接受外界光和温度等环境信号并传入核心振荡器,使得生物时钟与环境同步;核心时钟系统:自我维持的昼夜振荡器;输出系统:将生物钟产生的信号传递出去而控制生物行为和生理的节律变化。早期分子和遗传学研究主要关注昼夜节律振荡器的分子机制及神经生物学,阐明了昼夜生物钟节律的主要分子机制及相关神经网络。最近更多的研究关注生物钟信号是如何输入和输出。本文以果蝇运动节律的相关研究为主要内容,围绕生物钟输入系统、振荡器、输出系统这3个组成部分对昆虫生物钟研究进展进行总结。

关键词 昆虫, 昼夜节律, 生物钟, 钟基因, 神经网络

Advances in molecular regulation of insect circadian clock

CHEN Wen-Feng HE Chun-Xia AN Chun-Ju ZHAO Zhang-Wu^{**}

(1. Department of Entomology, China Agricultural University, Beijing 100193, China)

Abstract Insects display a wide range of circadian rhythms in their behavior and physiology, such as locomotor activity, sleep, learning and memory, mating and smell. Almost all animals have conserved circadian systems. Animal circadian systems include input pathways, a core clock system and output pathways. Input pathways are the means by which circadian oscillators receive information from the environment, such as light signals and temperature changes. This allows the oscillator to remain synchronized with the environmental day-night cycle. The core clock system is an oscillator that contains the molecular mechanism that generates self-sustained rhythmicity. Output pathways transmit signals from the oscillator to produce observable rhythms in behavior and physiology. Early genetic studies, which focused on molecular mechanisms of circadian oscillators and neurobiology, clarified the main molecular mechanism and neural networks involved in circadian rhythms. More recent research focuses on how the clock receives signals from the input system and transmits signals to the output system. In this paper, we summarize research on the insect circadian clock's input and output pathways and oscillator.

Key words insect, circadian rhythm, clock, clock genes, neural network

许多生物过程以内源时钟机制维持其自身的生理和行为过程以适应外界环境。从低等生物到高等生物,都处在昼夜交替的节律之中,它们的许多行为活动表现出相应的节律性。昆虫生物钟节律的研究是人类了解生物节律的重要途径。昆虫的许多生理和行为如睡眠、趋光性、学习记忆、羽化、取食和代谢、交配产卵、免疫等都表现出一定

的生物钟节律(Allada and Chung, 2010)。

生物钟由3个主要部分构成,1. 输入系统:接受外界环境信号并传入核心振荡器,使得生物时钟与环境同步;2. 核心时钟系统:是能够自我维持的昼夜振荡器,由一组呈现节律表达的钟基因和相应的蛋白组成,用于控制时间;3. 输出系统:将生物钟产生的信号传递出去而控制生物行为和

^{*} 资助项目:国家自然科学基金(31172090)。

^{**} 通讯作者, E-mail: zhaozw@cau.edu.cn

收稿日期:2011-10-06,接受日期:2011-10-25